



Data Hygiene: Good Data at the Ready



Reports Need Good Data. How many times have you run a report and realized that the data you had wasn't accurate? And you don't have time to deal with it because that report is needed ASAP for your annual campaign. So, rather than fix the data in Salesforce, you export the data, fix it in Excel, and say, "I'll get back to that later." But you never get back to it again because you are off to the next thing. Don't feel bad. EVERYONE has done this, but we know it's time to stop!

Having ready-to-use data is critical for user adoption as well as effective reporting. No one likes to find out at the end of year-end appeal season that essential data was wrong or missing all along! Yet, it's challenging to keep your data clean and ready-to-go when you have so many other things on your plate.

Where to Start? Having a plan in place to help with the data clean up can make life easier. A process of sorts that you can regularly run that outlines your data cleaning plan. There can be a lot of areas to hit, but we're going to narrow our focus on three main areas:

- Good Data
- Regular Checkup
- Freshening Data

It is possible to keep your data clean and make it easier when you have to run a report for an annual report or anything if you put together a plan and stick with it. You can begin this process of a plan by setting priorities and asking others to help you ensure data is accurate and clean. Creating reports and running campaigns will be as easy as a button click, and no more data cleaning a spreadsheet after the fact!

Good Data

Good data, what is that? It's vital to find a balance between keeping inaccurate data out of your system and keeping your users from using the system! Finding a place of comfort will vary with each organization and the users in that organization.

First, **developing a strategy around key fields** is vital. So, if you know you'll be reporting on some specific fields on an account, say household size or organization type. Then you want to ensure these fields are entered or make the field required. There are multiple ways to make a field required: (1) requiring the field on the object; (2) requiring the field on the page layout; (3) using a validation rule. Deciding which one to use will depend on your goals for the field (see the following chart).

Secondly, to **develop standards documentation** on how data should be entered for key reporting fields. For example, the Account name for an organization or business must contain their full corporate name as filed with the IRS. Or maybe your rule is that the Account name should not contain any extraneous information such as LLC or Corporation. These types of standards should be discussed as a team.

Finally, **duplicate management** can help stop or warn users before they create a potential duplicate. Salesforce.com has built-in duplicate management. It is possible to check before a Contact or Account is created to ensure it doesn't already exist. When thinking about using duplicate management, consider both the user and external systems (which is really another user). When considering a person who is a user, it is possible to give them a warning before they enter a potential duplicate. However, you won't want to set up a warning in an external system as it would require a decision. It may make more sense to allow the duplicate to be created but to have it reported as a possible duplicate to be cleaned later. Duplicate management is helpful, but be mindful of external systems.

The following chart explains strategies for having good data:

Name	Description	Recommendations
Required Fields	Forces the user to enter data into a field	<ul style="list-style-type: none"> • Identify the fields that are required to save a record <ul style="list-style-type: none"> ○ Begin with Accounts and Contacts ○ Step into other objects carefully • Requirement options: <ul style="list-style-type: none"> ○ <i>Object requirement</i> - won't allow a record to save if the field is not filled in. Use this one very carefully and thoughtfully, especially if using external tools to create Contacts/Accounts. ○ <i>Page Layout requirement</i> - at the object level, a field could still be left empty, but when doing data entry, the user will not be able to save the record. Stops bad data from being entered by users. ○ <i>Validation Rule requirement</i> - stops the user from entering invalid data after pressing the save button. This can allow for a bit more checking.
Field Entry Requirements	Forces the user to enter specific fields in a clearly defined method	<ul style="list-style-type: none"> • Identify the fields that must be filled in with specific requirements • Using Validation Rules <ul style="list-style-type: none"> ○ Is it possible to check that the field contains or does not contain specific information? • Using Process Builder or Workflow Field Updates <ul style="list-style-type: none"> ○ Can you build the field with existing fields on the record? For example, could you build an Opportunity Name with existing fields on the Opportunity?
Prevent Duplicates	Forces the users to review any record that are considered duplicates	<ul style="list-style-type: none"> • Identify what makes a record a duplicate. For example, on Contact, it could be the same first name, last name, and email is considered a duplicate <ul style="list-style-type: none"> ○ Would you want to stop users entirely from entering a duplicate, or could they create one after being notified? • Use conditions to allow duplicates to be created by specific users. For example, external systems should be allowed to create duplicates, so the data is captured.

Regular Checkups

Despite all your hard work, someone was able to enter bad data into the system. You will need to find that bad data and fix the information entered. As long as you know what to do. That's often the challenge.

First, how can you find this data? It can feel like a needle in a haystack. Still, you can do some specific things to figure out what is happening. Having a regular checkup of data is a perfect way to begin.

There are two things to consider when doing regular checkups: duplicates and data quality.

Duplicates

Finding duplicates can be the literal needle in a haystack. It can be even more challenging when you have a lot of data. Where to begin? Apsona offers Dedupe and Match a tool to help find that needle. It is a simple and Salesforce native tool that can help you find the duplicates in your existing instance and find matches from in CSV file or Contacts or Accounts.

Once you find the duplicates, then what? Suppose there are only a handful of duplicates. In that case, it is possible to do it using the Contact Merge tool that comes with the Nonprofit Success Pack, but what if you find HUNDREDS of them. The Apsona Dedupe tool provides a way to mass merge contacts, and it uses the NPSP Contact Merge functionality.

Apsona's Dedupe tool is quite powerful by helping you find the duplicates and then even helping with the merge process. But keep in mind that with power comes great responsibility. Before deduping, back up your data and/or do it in a sandbox that contains all your data from production.

Data Quality

How many times have you run a report of contacts and found that you have no actual contact information? Not much good if you cannot connect with them by some method. Data quality is ensuring the data that is entered has been filled in fully. Having a set of reports that you can efficiently run or use in a dashboard can help find these records with missing data.

Speaking of dashboards, there is a Data Quality Analysis Dashboard available for download on the AppExchange (see the Resources section below for the link). There are actually four different dashboards provided here. Still, the ones you might most interesting in are the High Level Data

Analysis and Account, Contact & Opportunity Data Quality Dashboards. Remember, these reports that run the Dashboards are canned. *HOWEVER*, you can modify these reports, thereby change the Dashboards to be focused on your data requirements.

The following chart gives you a suggested list of reports to create to check your data quality. Running them on a schedule to be sent to you regularly isn't a bad idea to keep your head in the data quality game. The NPSP also provides a list of **Health Check reports**, and speaking of the NPSP, **running the Health Check** isn't a bad idea either! It does check your data quality.

Suggested Data Quality Reports

Name	Description
Contacts missing Contact Data	You define the fields you wish to see filled in, but this report should be a list of contacts with data missing. So, using filtering criteria is vital, such as Email equals {blank} OR Phone equals {blank}, etc.
(NPSP Specific) Contact without Preferred Phone Number and/or Email Address	This report looks specifically at the Preferred Phone Number & Preferred Email Address fields to ensure it is selected. Mostly this is selected by default, but sometimes it does not get filled in.
Accounts without Contacts	Use this report to find Accounts without Contacts (as a hint, use Cross Filters on an Account report).
Object summarized by record owner (Contacts, Accounts, Leads, Opportunities, and other objects used)	Create a bunch of reports summarized by Record Owner and ensure that all the Owners are active users of Salesforce. If they are not, then change the ownership to an active user.

Regular Checkup Plan

Use this chart to help in creating a plan for a regular checkup with your data. Fill in the "**how often?**" for your own organization depending on how much your data is used

What to Do?	How Often?	What to Look For?
Finding Duplicates		<ul style="list-style-type: none">• Can you tell how the duplicates were created?• How many duplicates are there in the system?• Is it really a duplicate or just one that appears to be one?
Merging Duplicates		<ul style="list-style-type: none">• Can you systematically find the winning record?• What to do if both records have a different email address, phone numbers, addresses, etc.?
Data Quality Reports		<ul style="list-style-type: none">• Who can find the missing data?• If a contact has no contact data, do they exist?• Who should inactive user records go to?• Is it okay to have an Account without any Contacts?
Dealing with Quality Issues		<ul style="list-style-type: none">• What is causing the contacts to lose their contact information?• Should I move records to active users after a user leaves the organization?

Fresh Data

Keeping your data fresh means caring for your existing data, Contacts, Accounts, Opportunities, and other objects in use. Managing address changes, phone number updates, and email changes can be a full-time job in itself. Picking a few data points you wish to keep updated can help you stay focused. Using tools that can help keep the information updated automatically is even better (but these are often not free).

First, deciding which fields you wish to keep up to date, it likely won't be all fields or records for that matter. Maybe there are a few Contacts and/or Accounts you wish to keep fresh. Perhaps it all of them. Deciding will make it easier when looking for the right tool.

Once you have picked fields and/or records, look for an AppExchange application that could be used. There are tools out there to scrub data, such as addresses and phone numbers. Perhaps with email addresses, you need to keep your data synced with your external email service. Or maybe another email tool that captures bounced and updated email addresses. It could also mean you need to update the information manually or regularly. Either via a phone call, email survey, or perhaps even a mailed letter.

Getting the entire staff involved in a "data clean up day" can also bring home the importance of good, quality data. So often, we know and care, but others don't seem to understand. Having them "feel the pain" of keeping the data clean can help you with the adoption of your processes and procedures.

Having fresh data you can trust is vital to a well-used and adopted system.

Resources

Here is a list of resource to help you with your data hygiene journey:

- [Manage Duplicate Records](#) – helps with setting up managing duplicates in Salesforce.
- [Data Quality Trail](#) - one hour trail that helps you with the issues of data quality. There are some excellent recommendations on this trail.
- [Apsona Dedupe and Match](#) – they offer a [nonprofit discount](#) for this product.
- [Data Quality Analysis Dashboard](#) – works on standard fields on most standard objects.



Document and Celebrate!

After you've gotten over this mountain, we hope you share the results with your team and community on the Power of Us Hub. We'd also love to hear about how it went in tightening with data hygiene, especially if you find a report that you think could help others. It's fun to celebrate together.